



TITLE:

スーパーコンピューターによる物理学：素粒子論から蛋白工学まで

AUTHOR(S):

岡本, 祐幸

CITATION:

岡本, 祐幸. スーパーコンピューターによる物理学：素粒子論から蛋白工学まで. 物性研究 1994, 62(2): 231-248

ISSUE DATE:

1994-05-20

URL:

<http://hdl.handle.net/2433/95341>

RIGHT:

スーパーコンピューターによる物理学:^{*} 素粒子論から蛋白工学まで

かうじ

奈良女子大学理学部 岡本祐幸

(1994年3月8日受理)

1. はじめに

最近のスーパーコンピューター技術の進歩にはめざましいものがある。それは物理学に限らず、多くの研究分野でこれまで不可能と思われてきた計算をどんどん可能にしている。ここで興味深いのは、全く別の分野の研究がスーパーコンピューターという同じ道具を使用すると共に、その計算に使われるアルゴリズム（例えばモンテカルロ法や分子動力学法）が多くの場合共通であることである。そこには、自然と極めて学際的な環境が想定できる。つまり、違う分野の研究者がスーパーコンピューターという同じ土俵で話ができ、例えば、ひとつの分野で開発された新しいアルゴリズムを他の分野に「輸出」することによって、その分野の研究の大幅な発展を促すことも可能になる訳である。実際、アメリカでは、このことに目をつけた人々が、政府や企業に働きかけて、既に10年程前に各地に共同利用のスーパーコンピューター研究所を設立した。例えば、コーネル大学には、Cornell National Supercomputer Facility が、フロリダ州立大学には、Supercomputer Computations Research Institute (SCRI) が、また、ミネソタ大学には、Supercomputer Institute 等5つ以上の研究所が、相前後して設立された。そして、それぞれの研究所の構成メンバーは、理論系及び実験系の物理屋が中心であるが、物理学以外にも、化学、生物、応用数学、地学、海洋学、機械工学等から、経済の分野に至るまで、実に多岐にわたり学際的である。ハードウェアとしては、例えば、Minnesota Supercomputer Institute の場合、Cray-2、Cray X-MP、Cray Y-MP、更にはCM-200、CM-5、及びT3D（導入予定）の使用が可能で充実している。

ところで、筆者がコーネル大学大学院に在学中の1982年、所属研究室の教授 K. Wilson がくり込み群でノーベル物理学賞を受賞するという慶事があった。受賞後、Wilson が急にネクタイをして背広を着だしたのであるが、それは企業や政府の人に頻繁に会って、スーパーコンピューター研究所設立の働きかけをしているからだと学生の間には噂になった。学内でも講演をして、「このままでは、アメリカのスーパーコンピューター産業が日本との競争に負けてしまう。」と熱弁していた。少し大げさに聞こえるが、案外こういう「殺し文句」が予算決定に重要だったようで、例えば、SCRI紹介のパンフレットにも、次のような文章が繰り返し出てくる。

^{*} 本稿は、編集部の方から特にお願ひして執筆していただいた記事である。

“At its inception, SCRI became symbolic of federal efforts to place the power of supercomputers in the hands of more scientists and students, answering a national imperative to protect the country’s lead in science and technology, grown fragile through intensified foreign competition.(中略) With near-daily news of breakthroughs in the design, performance, and versatility of American-made supercomputers, SCRI enters a new decade with a sense of pride in having played a major role in strengthening the nation’s position on the world market of scientific innovation through supercomputing.”

その「競争相手」の日本の方はどうかというと、残念ながら、アメリカの研究所に匹敵する程のものは、10年以上たった現在でもできていない。今後、大規模なスーパーコンピュータ研究所が各地にできて、より多くの研究者に十分な計算時間を提供できるようになることを期待したい。

さて、本稿では、筆者が直接関わった、コンピューターを利用した研究について、簡単な解説を試みることにする。2. ではレプトンの異常磁気能率、3. では調和振動子の力学系を用いた大エルミート行列の対角化法、4.1. では徐冷モンテカルロ法による、また、4.2. ではマルチカノニカル法による蛋白質の立体構造予測について述べる。

2. レプトンの異常磁気能率

物理学に於て、理論値と実験値が最も合致している物理量のひとつは電子の異常磁気能率であろう。以下に述べるように、それは8桁目まで一致している。その理論計算の複雑さは並大抵ではなく、実際、8次の摂動計算を最後までやり遂げたのは、コーネル大学の木下グループ唯ひとつである（詳しくは文献1参照）。その計算はスーパーコンピューターの能力を最大限に利用したものであり、この数年の間に初めて精度の高い計算ができるようになったと言える。

電子やミューオンのように、強い相互作用をしない素粒子をレプトンと呼ぶが、磁場中のレプトンの有効ハミルトニアンは次で与えられる。

$$H_{eff} = -g\mu_B \vec{S} \cdot \vec{B} \quad (1)$$

ここで、 g は g 因子、 μ_B はBohr磁子、 \vec{S} はレプトンのスピン、 \vec{B} は磁束密度である。Diracの相対論的量子力学の成功のひとつは、古典的に厳密に $g = 2$ が成立ち、実験値と良い一致が得られたことである。ところが、実験技術が進むにつれ、 g が2から僅かにずれていることが発見され(異常磁気能率)、Schwingerがそのずれが量子効果であると直感し、2次の摂動計算で見事に説明した²⁾ことはQED(量子電磁力学)黎明期の記念すべき出来

事であった。異常磁気能率は次のように摂動展開できる。

$$a \equiv \frac{1}{2}(g - 2) = C_1 \frac{\alpha}{\pi} + C_2 \left(\frac{\alpha}{\pi} \right)^2 + C_3 \left(\frac{\alpha}{\pi} \right)^3 + \dots \quad (2)$$

ここで α は微細構造定数である。Schwinger の結果は $C_1 = \frac{1}{2}$ であり、対応する Feynman 図は 1 個である (図 1 参照)。

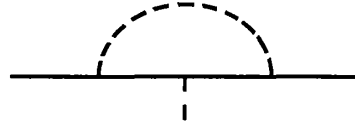


図 1. 異常磁気能率の 2 次の摂動計算に対応する Feynman 図。実線はレプトン、点線は光子を表す。

異常磁気能率が QED の最も厳密な検証の手段と成り得たのは、その発見以来現在に至るまで実験値及び理論値の両方の精度が次から次へと改善されてきたからである。電子の異常磁気能率 a_e の摂動計算に於いては、4 次には 7 個、6 次には 72 個、8 次には 891 個の Feynman 図がそれぞれ存在する。現在では、6 次までなら大部分の寄与が解析的に求まっている。しかし、8 次の項 (C_4) への寄与はほとんど数値積分に頼らざるを得ず、ここにスーパーコンピュータが必要となるのである。(被積分関数は大きいもので 2 万項にも及び、積分次元は 10 次元である。) 木下等の 10 年以上にわたる精力的な計算の結果は次の通りである³⁾。

$$C_4 = -1.434(138) \quad (3)$$

ここで括弧内の数字は数値積分の誤差を表す。これに低次の寄与

$$\begin{aligned} C_1 &= 0.5 \\ C_2 &= -0.328\,478\,965\dots \\ C_3 &= 1.176\,11(42) \end{aligned} \quad (4)$$

及びハドロンや弱い相互作用の寄与等 (誤差内で無視できる量) を足し、量子ホール効果から求められた α の値⁴⁾

$$\alpha^{-1} = 137.035\,997\,9(32) \quad (5)$$

を代入すると (式 (2) 参照) 電子の異常磁気能率は以下の値になる³⁾。

$$a_e(\text{theory}) = 1\,159\,652\,140(28) \times 10^{-12} \quad (6)$$

ここで誤差の大部分は (5) の α の測定誤差である。一方、最新の実験値は

$$\begin{aligned} a_{e-} &= 1\,159\,652\,188.4(4.3) \times 10^{-12} \\ a_{e+} &= 1\,159\,652\,187.9(4.3) \times 10^{-12} \end{aligned} \quad (7)$$

であり⁵⁾、理論値と実験値は 8 桁まで一致している。逆に、理論値と実験値を等しく置くことによって α を求めると

$$\alpha^{-1}(a_e) = 137.035\,992\,22(94) \quad (8)$$

となり、(5) と比べても分かるように最も精密な α の値を得られることになる³⁾。

電子の異常磁気能率はほとんどが QED の寄与で占められているのに対し、ミューオンの異常磁気能率 a_μ では弱い相互作用や強い相互作用からの寄与も重要になってくる。これはミューオンの質量が電子のその約 200 倍と重いことに起因する。強い相互作用の a_μ への寄与は次の通りである⁶⁾。

$$a_\mu^{had} = 7027(175) \times 10^{-11} \quad (9)$$

また、弱い相互作用の a_μ への最低次の寄与は

$$a_\mu^{weak} = 195(10) \times 10^{-11} \quad (10)$$

で与えられる⁷⁾。更に、8 次の QED の寄与は 1360 個の Feynman 図から成り立つがその内の a_e との共通部分 ((3) 参照) 以外の寄与は次のように求められた⁸⁾。

$$C_4^\mu - C_4^e = 127.00(41) \quad (11)$$

図 2 にこの寄与に対応する Feynman 図の典型的な例を幾つか示した。さて、10 次の QED の寄与の概算値

$$C_5^\mu = 570(140) \quad (12)$$

等も含めた全寄与を足すと次の理論値が得られる⁸⁾。

$$a_\mu(\text{theory}) = 11\,659\,192(18) \times 10^{-10} \quad (13)$$

一方、実験値は

$$\begin{aligned} a_{\mu-} &= 11\,659\,370(120) \times 10^{-10} \\ a_{\mu+} &= 11\,659\,110(110) \times 10^{-10} \end{aligned} \quad (14)$$

であり⁹⁾、良い一致を示している。

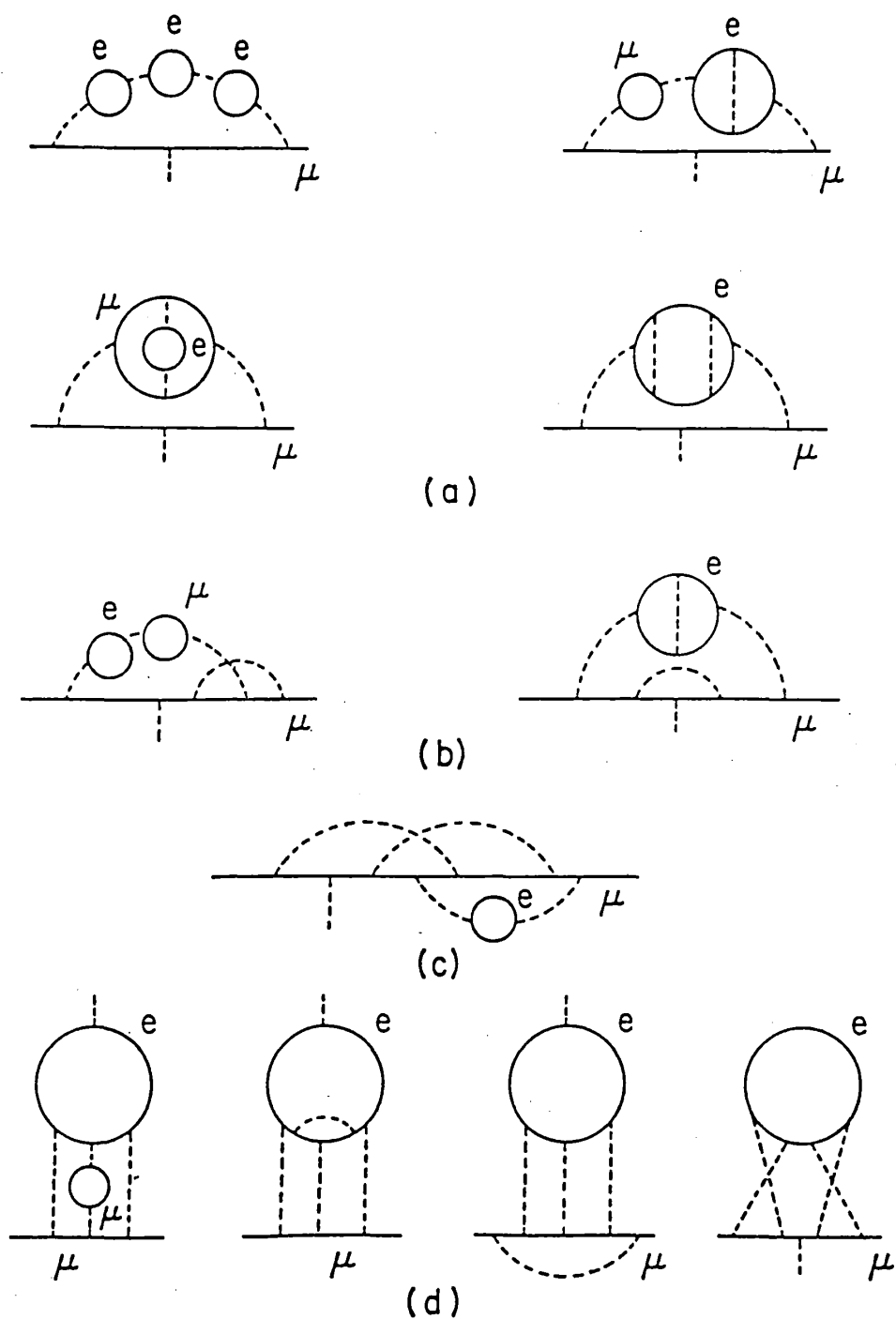


図 2. 異常磁気能率の 8 次の摂動計算に対応する Feynman 図の代表的なもの。

ここでは、物理学に於て理論値と実験値が最も精度良く一致している例を示した。異常磁気能率の計算は QED のくり込みが高次の摂動まで矛盾なくできることを示す例にもなっている。ミューオンの異常磁気能率では理論値の誤差が弱い相互作用の寄与とほぼ同じにまで小さくなっており、実験値の誤差が改善される時、Glashaw-Weinberg-Salam 理論を 1 ループのオーダーで検証する 1 例となることを意味する。このことをふまえて、Brookhaven 国立研究所で新しいミューオン異常磁気能率の測定実験が始まっている¹⁾。また、我々の a_μ の 8 次の計算⁸⁾とくり込み群による概算値を比べることから、くり込み群方程式にでてくる、Callan-Symanzik の β 関数の QED の高次の項の従来の計算法が 6 次までは正しいが、8 次以降は補正項を計算する必要があることが発見された。¹⁰⁾

3. 大エルミート行列の対角化法

物理学に限らず、いろいろな分野の研究に於て、大エルミート行列の最大または最小固有値及び固有ベクトルの計算が要求されることがよくある。しかし、行列の次元が大きいと、必要な計算時間及びメモリー容量が膨大になってしまって、従来の組み込みルーチンでは手に負えなくなってしまう。現在のところ、最も広く使われている方法のひとつが、Lanczos 法であろう¹¹⁾。ここでは、Lanczos 法と同じオーダーの計算時間及び容量を必要とするが、極めて簡単な原理に基づく新しい対角化法を紹介する¹²⁾。この方法は調和振動子の力学系の共鳴現象を用いて、系の状態密度を出す方法^{13,14)}から導かれた。

対角化したい $N \times N$ エルミート行列を $\phi = (\phi_{ll'})$ とする。この問題は、 N 個の粒子の格子の基準振動のモードを求めることと同等である。つまり、行列要素 $\phi_{ll'}$ を振動子 l と l' を結ぶバネのバネ定数とみなして、次の運動方程式を考える。

$$M\ddot{u}_l(t) = - \sum_{l'=1}^N \phi_{ll'} u_{l'}(t), \quad (l = 1, \dots, N) \quad (15)$$

ここで、 M と $u_l(t)$ はそれぞれ l 番目の粒子の質量と変位である。以下簡単のため、 ϕ を実対称行列とし、 $M = 1$ とおく。(一般のエルミート行列の場合への拡張は自明である。) 時間がステップ τ で離散化されたとし、微分を差分にして、 τ の 1 次までの近似で (15) は次のように書き換えられる。¹³⁾

$$\begin{aligned} v_l(n+1) &= v_l(n) - \tau \sum_{l'=1}^N \phi_{ll'} u_{l'}(n) \\ u_l(n+1) &= u_l(n) + \tau v_l(n+1) \end{aligned} \quad (16)$$

ここで、 $v_l(n)$ は l 番目の粒子の時刻 $t = n\tau$ での速さである。また、 $u_l(n)$ と $v_l(n)$ は基準

振動モードの和として次のように書ける。

$$\begin{aligned} u_l(n) &= \sum_{\lambda} Q_{\lambda}(n) e_l(\lambda) \\ v_l(n) &= \sum_{\lambda} P_{\lambda}(n) e_l(\lambda) \end{aligned} \quad (17)$$

ここで、 $Q_{\lambda}(n)$ と $P_{\lambda}(n)$ はモード λ がそれぞれ $u_l(n)$ 及び $v_l(n)$ へ寄与する振幅であり、 $e_l(\lambda)$ はモード λ の規格直交化された固有ベクトルである。

$$\begin{aligned} \sum_{l'=1}^N \phi_{ll'} e_{l'}(\lambda) &= \mu_{\lambda} e_l(\lambda) \\ \sum_{l=1}^N e_l(\lambda) e_l(\lambda') &= \delta_{\lambda\lambda'} \end{aligned} \quad (18)$$

ここで、 μ_{λ} は行列 ϕ の固有値である。

さて、(17) を (16) に代入し、(18) を使い、更に $P_{\lambda}(n)$ を消去すると、次が得られる。

$$Q_{\lambda}(n+1) - (2 - \mu_{\lambda}\tau^2)Q_{\lambda}(n) + Q_{\lambda}(n-1) = 0 \quad (19)$$

よって、 $Q_{\lambda}(n) = (\beta_{\lambda})^n$ (β_{λ} は定数) という解を仮定すると、次が得られる。

$$\beta_{\lambda} = \frac{2 - \mu_{\lambda}\tau^2 \pm \sqrt{\mu_{\lambda}\tau^2(\mu_{\lambda}\tau^2 - 4)}}{2} \quad (20)$$

この解から分かることは、 τ (及び μ_{λ}) の大きさによって、次の二つの場合があることである。

$$\begin{aligned} (A) \quad & |\beta_{\lambda}| = 1 \quad \text{for} \quad 0 \leq \mu_{\lambda}\tau^2 \leq 4 \\ (B) \quad & |\beta_{\lambda}| \neq 1 \quad \text{for} \quad \mu_{\lambda}\tau^2 < 0 \quad \text{or} \quad \mu_{\lambda}\tau^2 > 4 \end{aligned} \quad (21)$$

特に、(B) に於いては、一つの根が $|\beta_{\lambda}| > 1$ を満たし、もう一つの根が $|\beta_{\lambda}| < 1$ を満たす。

よって、(A) の時、 Q_{λ} は振動し、同じオーダーの大きさに留まるが、(B) の時、 Q_{λ} は時間と共に指数関数的に大きくなっていく。本方法は正にこの性質を利用する。ここで、注意しておきたいことは、同じ固有値 μ_{λ} でも、時間ステップ τ の値によって、(A) の場合になったり、(B) の場合になったりすることである。

以下、簡単のため、全ての固有値が 0 以上であるとする。すると、最大固有値 $\mu_{\lambda_{max}}$ は次のように求められる。 $\mu_{\lambda_{max}}$ については、 $\mu_{\lambda}\tau^2 > 4$ であり、残りの固有値 μ_{λ} については、 $0 \leq \mu_{\lambda}\tau^2 \leq 4$ となるように、時間ステップ τ を選ぶ。そして、任意の初期条件から、(16) の時間発展をさせると、(21) より、最大固有値に対応するモードの振幅 $Q_{\lambda_{max}}$ のみが指数関数的に大きくなっていく。つまり、 $u_l(n)$ は時間と共に、 $\mu_{\lambda_{max}}$ に対応する固有ベ

クトル $e_l(\lambda_{max})$ の定数倍に収束する。また、固有値 $\mu_{\lambda_{max}}$ は最終の $u_l(n)$ を使って、

$$\mu_{\lambda_{max}} = \frac{\sum_{l,l'=1}^N u_l(n) \phi_{ll'} u_{l'}(n)}{\sum_{l=1}^N u_l(n)^2} \quad (22)$$

で与えられる。次の固有値及び固有ベクトルは、上の手順を $e_l(\lambda_{max})$ と直交する部分空間に限定することによって、順次求めていくことができる。

さて、最適の τ の値は、系の全ポテンシャルエネルギーを計算することによって、簡単に求められるが、その詳細は文献 12 にゆだねることにして、1 次元 $s = \frac{1}{2}$ 反強磁性 Heisenberg モデルに於ける計算例を以下に記す。 N_S 個のスピンの系のハミルトニアンは

$$H = \sum_{i=1}^{N_S} \vec{S}(i) \cdot \vec{S}(i+1) \quad (23)$$

で与えられる。ここで、 $\vec{S}(i)$ は格子点 i のスピン演算子である。また、周期境界条件を課す。このハミルトニアン行列の次元は $N = 2^{N_S}$ であり、 $N_S = 14$ (i.e., $N = 16384$) までの場合の系の基底状態及び第一励起状態のエネルギー固有値を計算してみた¹²⁾。表 1 に本アルゴリズムで固有値が時間のイテレーションと共に、どのように収束していくかを示した。少なくともこのモデルに於いては、収束に必要なイテレーション回数は行列の大きさを変えても、あまり変わらない。(60 回程度で十分である。) これらの固有値は、Lanczos 法の結果¹⁵⁾と 7 桁までの一致を見た。

最後に、この方法は高度のベクトル化が可能なので、スーパーコンピュータでの計算に適していることを付記する。

表 1. Heisenberg 反強磁性鎖の基底状態と第一励起状態のエネルギー (μ_{λ_0} と μ_{λ_1})。

IT	$\mu_{\lambda_0}(N = 64)$	$\mu_{\lambda_0}(N = 1024)$	$\mu_{\lambda_0}(N = 16384)$
20	-2.8027750983	-4.5153136281	-6.2632192228
40	-2.8027756377	-4.5154463545	-6.2635493911
60	-2.8027756377	-4.5154463545	-6.2635495335
80	-2.8027756377	-4.5154463545	-6.2635495335
100	-2.8027756377	-4.5154463545	-6.2635495335
IT	$\mu_{\lambda_1}(N = 64)$	$\mu_{\lambda_1}(N = 1024)$	$\mu_{\lambda_1}(N = 16384)$
20	-2.1180325187	-4.0922058679	-5.9564352761
40	-2.1180339887	-4.0922073467	-5.9564438234
60	-2.1180339887	-4.0922073467	-5.9564438240
80	-2.1180339887	-4.0922073467	-5.9564438240
100	-2.1180339887	-4.0922073467	-5.9564438240

4. 第一原理からの蛋白質の立体構造予測

蛋白質は自然環境下で一定の形状（特異立体構造）を持っている。それぞれの蛋白質の特異立体構造はその生化学的機能と深い関わりを持つため、理論的に立体構造を予測することは生物物理学に於ける重要な問題のひとつである。蛋白質の立体構造はそのアミノ酸配列の情報のみで決まっていると信じられているが、未だにその情報のみから *ab initio* 的に立体構造を予測できた例はない。この問題の困難点は、系に局所的極小エネルギー状態が無数に存在するため、シミュレーションがそのどこかに留まってしまって、なかなか全位相空間をカバーできないことにある。よって、計算結果が初期状態に強く依存する事になる。具体的に、分子動力学法を例にとると、現在世界一高速のスーパーコンピュータをもってしても、小さな蛋白質に於て、ナノ秒程度のシミュレーションが現実的に計算可能なタイムスケールの上限である。ところが、蛋白質の折りたたみの（実験的）タイムスケールはミリ秒またはそれ以上であるから、これは、大雑把に言って、コンピュータの速度が最低 10^6 倍速くなしないと、上に述べた *ab initio* な立体構造予測は不可能であることを意味する。よって、何らかの有効なアルゴリズムの開発が切望されている。

蛋白質の立体構造予測の困難は、実は、フラストレーションを持つ系（例えば、巡回セールスマン問題、スピングラス、ニューラルネットワーク、etc.）に共通の難問で、所謂 NP complete な最適化問題として有名である。よって、他の分野で開発された有効なアルゴリズムを導入することが考えられる。実際、筆者らは、徐冷モンテカルロ法（Monte Carlo simulated annealing）¹⁶⁾ 及びマルチカノニカル法（multicanonical algorithm）¹⁷⁾ を蛋白質の立体構造予測問題に適用することを提唱してきた。

本研究の計算に於いては、蛋白質の系のエネルギーとして、静電エネルギー、Lennard-Jones エネルギー、水素結合エネルギーの各原子対に関する和に、ねじれエネルギーの各回転角に関する和を加えたものを用いた。

$$\begin{aligned}
 E_{total} &= E_{el.} + E_{L.J.} + E_{h.b.} + E_{tors.} \\
 E_{el.} &= \sum_{(i,j)} \frac{332 q_i q_j}{\epsilon r_{ij}} \\
 E_{L.J.} &= \sum_{(i,j)} \left(\frac{F_{ij} A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) \\
 E_{h.b.} &= \sum_{(i,j)} \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) \\
 E_{tors.} &= \sum_i U_i (1 \pm \cos(n_i \chi^i))
 \end{aligned} \tag{24}$$

ここで、 r_{ij} は原子 i と原子 j の距離 (Å)、 ϵ は誘電率を表す。また、 $E_{el.}$ 中の係数 332 は

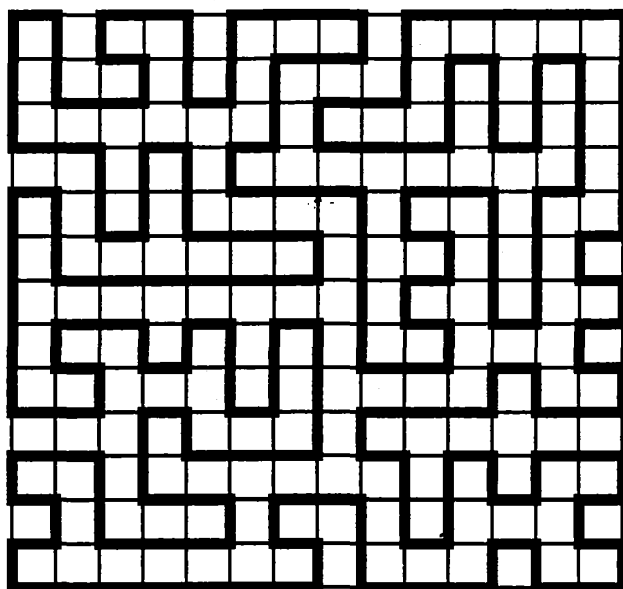
エネルギーを kcal/mol の単位で表すための定数である。上式にでてくる種々のパラメーター及びアミノ酸の標準分子構造は ECEPP/2¹⁸⁾ のものを採用した。また、溶媒の効果はほとんどの場合、誘電率 $\epsilon = 2$ という形で、平均場的に取り入れた。ペプチド結合の二面角 ω は簡単のため 180° に固定し、残った主鎖の全ての二面角 ϕ と ψ 及び側鎖中の全ての二面角 χ を自由度とした。よって、アミノ酸数が増えると、自由度の数が大きくなって、非常に複雑な系となることが予想できよう。

次の節では、まず、徐冷モンテカルロ法を、その次の節では、マルチカノニカル法を用いた構造予測について述べる。

4.1. 徐冷モンテカルロ法による立体構造予測

徐冷モンテカルロ法は物質の結晶を作る時のプロセスに基づいている。つまり、十分高温でシミュレーションを始め、徐々に温度を下げながらエネルギー極小値に留まることなく、最終的に最小エネルギー状態（結晶）に到達しようというものである。温度の下げ方が熱平衡を保つ程緩やかならば、最終温度を 0 K とする時に、ポテンシャルエネルギーの最小状態のみが得られる。また、最終温度をそれ以上の、ある有限の値とするならば、得られる最終状態は一意的ではなく、その温度でのカノニカル分布を近似することになる。この方法の利点は従来の温度一定のシミュレーションと違って、エネルギーの極小状態に留まることが避けられるため、計算結果が初期状態に依存しなくなり、また熱平衡への到達も大幅に加速される事である。

図 3. 都市数 210 の巡回セールスマン問題の徐冷モンテカルロ法による解の例。



徐冷法の有効性をイラストレートするために、ここで、都市数 210 の巡回セールスマン問題に於ける結果を示す。この場合、可能な経路の数は $209! \approx 10^{400}$ と膨大で、最短経路を捜す問題の難しさが実感できよう。正解が分かりやすいように、町を格子上に分布させて、徐冷モンテカルロ法を適用して得られた解の一つを図3に示す。この結果は、奈良女子大学のスカラーマシーン FACOM M770/6 で約 30 分の CPU 時間で得られたものである。

さて、徐冷モンテカルロ法の蛋白質の *ab initio* 立体構造予測への適用は数年前に提唱された。¹⁹⁾ 以来、筆者等の研究グループは、主に小さなペプチドに於てこの方法の有効性を確かめてきた。具体的な手法は次の通りである。完全にランダムな構造を初期状態とする。各モンテカルロステップに於ては、二面角を順次一回ずつ update し、その度に Metropolis 判定を課す。勿論、update は完全にランダムに行い、特定の終状態に向かうような重みは一切付けていない。徐冷の仕方としては、初期温度及び最終温度を各々 1000 K 及び 250 K とし、モンテカルロステップと共に指数関数的に温度を減少させる。1 回のシミュレーションは 10^4 回のモンテカルロステップから成り、ひとつのペプチドについてシミュレーション数は 20 回とする。以上の手法に基づき、様々なペプチドに於て徐冷モンテカルロシミュレーションを実行した。まず、リボヌクレアーゼ A の C-ペプチドでは X 線構造解析で示唆されている位置に 2.5 周の α -ヘリックスが形成される事を示すと共に、グルタミン酸 Glu-2⁻ とヒスチジン His-12⁺ の側鎖の電荷による静電相互作用がこの α -ヘリックス構造の安定性に重要であるという CD 実験の結果を検証した。²⁰⁾ 次に、ウシ膵臓トリプシンインヒビターのフラグメント BPTI(16-36) では α -ヘリックスは 1 回もできない代わりに、X 線解析で示唆されている位置の近傍に β -シート構造ができる事を示した。²¹⁾ また、副甲状腺ホルモンのフラグメント PTH(1-34) に於ては、NMR 実験で示唆された場所 2 箇所にも α -ヘリックスができることも示した。²²⁾ 更に、アミノ酸のホモポリマーに於て、 α -ヘリックス形成傾向性を調べ、それが蛋白質の X 線解析のデータベースからの示唆と良く合うことを示した。^{23,24)} そして、17 残基の合成ペプチドに於いて、得られた α -ヘリックス構造が側鎖も含めて、NMR 及び CD 実験から示唆される構造を良く再現していることを示した。²⁵⁾ 参考のため、この合成ペプチドに於いて、徐冷モンテカルロ法が予測した α -ヘリックス構造のステレオ図を図4に示した。

以上の結果は上述の全く同じ手法に基づき、完全にランダムな初期構造からシミュレーションを始めて得られたものであり、アミノ酸配列の違いが直接結果の違いに反映される事を実証している。

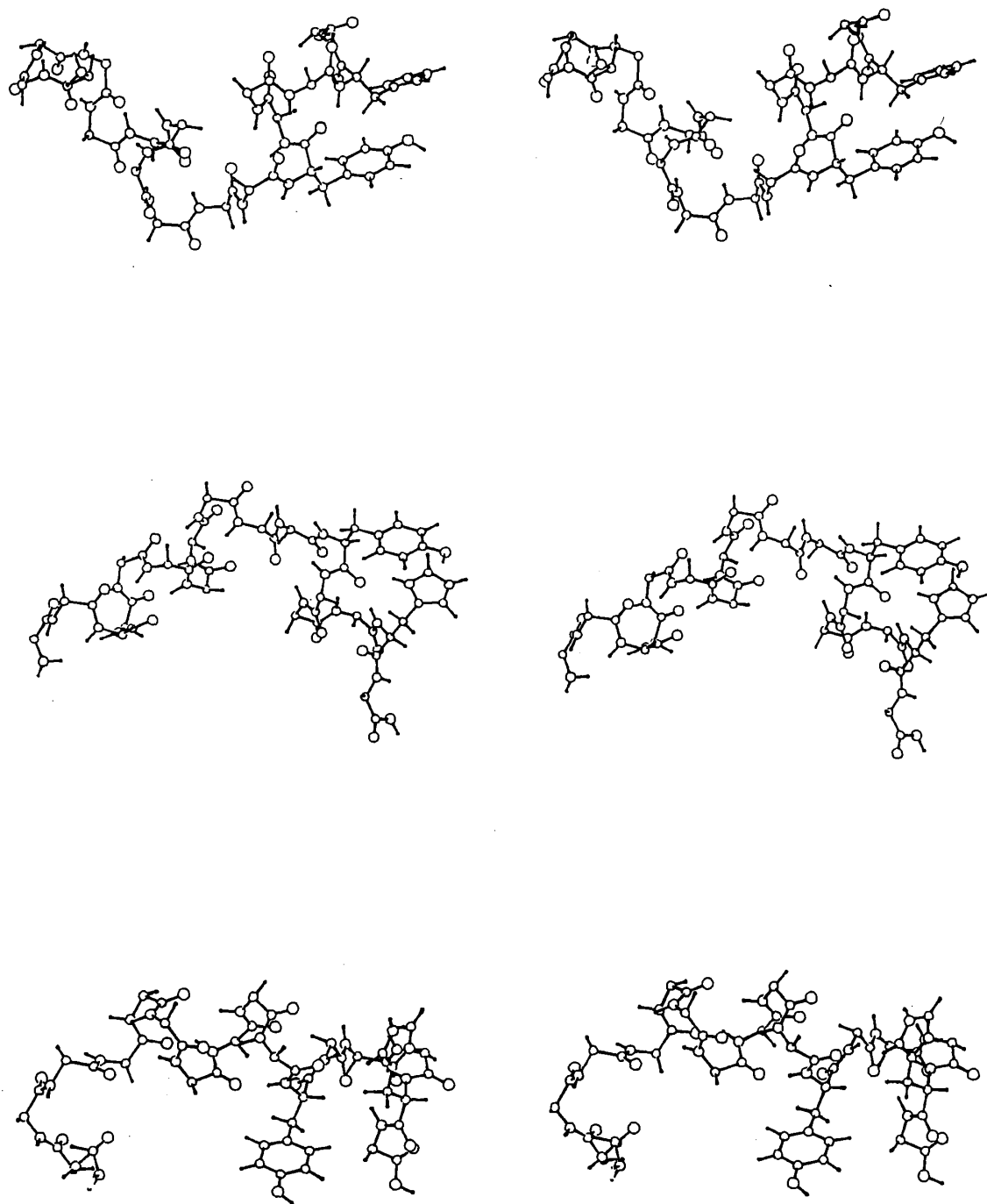


図 4. 徐冷モンテカルロ法で予測された 17 残基の合成ペプチドの構造のステレオ図。

4.2. マルチカノニカル法による立体構造予測

マルチカノニカル法は、普通のモンテカルロ法がカノニカル分布に従って状態を発生させるのに対し、どのエネルギー状態も同じ重みで（「マルチカノニカル分布」に従って）発生させ、エネルギー空間上の1次元ランダムウォークを実現する方法である。よって、どんなエネルギーの障壁も乗り越えられ、極小値に留まるのを避けられる。マルチカノニカル法の利点は唯一回のシミュレーションを実行するだけで、最小エネルギー状態ばかりでなく、任意の温度におけるカノニカル分布が求められる事である。

以下、この方法を少し詳しく説明しよう。普通のカノニカル集合に於いては、温度の逆数 $\hat{\beta} \equiv 1/RT$ が与えられた時、それぞれの状態は次の Boltzmann 因子の重みを持つ。

$$W_B(E) = \exp(-\hat{\beta}E) \quad (25)$$

すると、エネルギーについての確率分布は次で与えられる。

$$P_B(\hat{\beta}, E) \propto n(E)W_B(E) \quad (26)$$

ここで、 $n(E)$ はエネルギー E を持つ状態の数である。状態数 $n(E)$ は増加関数であるが、Boltzmann 因子は E と共に指数関数的に減少するので、 $P_B(\hat{\beta}, E)$ は一般にベル型の形をしている。これに対し、「マルチカノニカル集合」に於いては、確率分布はどのエネルギーの状態も同じ確率で存在するように定義される。つまり、次が成り立つ。

$$P_{mu}(E) \propto n(E)W_{mu}(E) = \text{const.} \quad (27)$$

すると、マルチカノニカル集合の重み因子は次を満たす。

$$W_{mu}(E) \propto n^{-1}(E). \quad (28)$$

この重み因子の具体的な形を決めるために、次の二つのパラメーター $\alpha(E)$ と $\beta(E)$ を導入する。¹⁷⁾

$$W_{mu}(E) \equiv e^{-B(E)} = \exp\{-\beta(E)E - \alpha(E)\}. \quad (29)$$

これらのパラメーター $\alpha(E)$ と $\beta(E)$ の求め方の詳細は文献 26 にゆだねる。 $\alpha(E)$ と $\beta(E)$ が求まったら、一回だけ長いマルチカノニカルシミュレーション（production run）を実行する。そして、この production run から最小エネルギー状態ばかりでなく、任意の温度 $\hat{\beta}^{-1}$ におけるカノニカル分布が求められる訳である。後者は re-weighting 法²⁷⁾を使って、次の式によって得られる。

$$P_B(\hat{\beta}, E) = \frac{e^{B(E)-\hat{\beta}E} P_{mu}(E)}{\sum_E e^{B(E)-\hat{\beta}E} P_{mu}(E)}. \quad (30)$$

さて、マルチカノニカル法の有効性をイラストレートするために、ここで、2次元 10×10 格子上の Ising モデルに於ける結果を示す。簡単のため、モンテカルロステップ数 10^4 回の短い production run を一回実行した。図5に内部エネルギー U と比熱 C の温度変化を与えた。星印はそれぞれの温度に於けるカノニカルモンテカルロシミュレーションの結果である。一回のマルチカノニカル production run の結果（実線）がそれぞれの温度の値を良く再現していることが分かる。

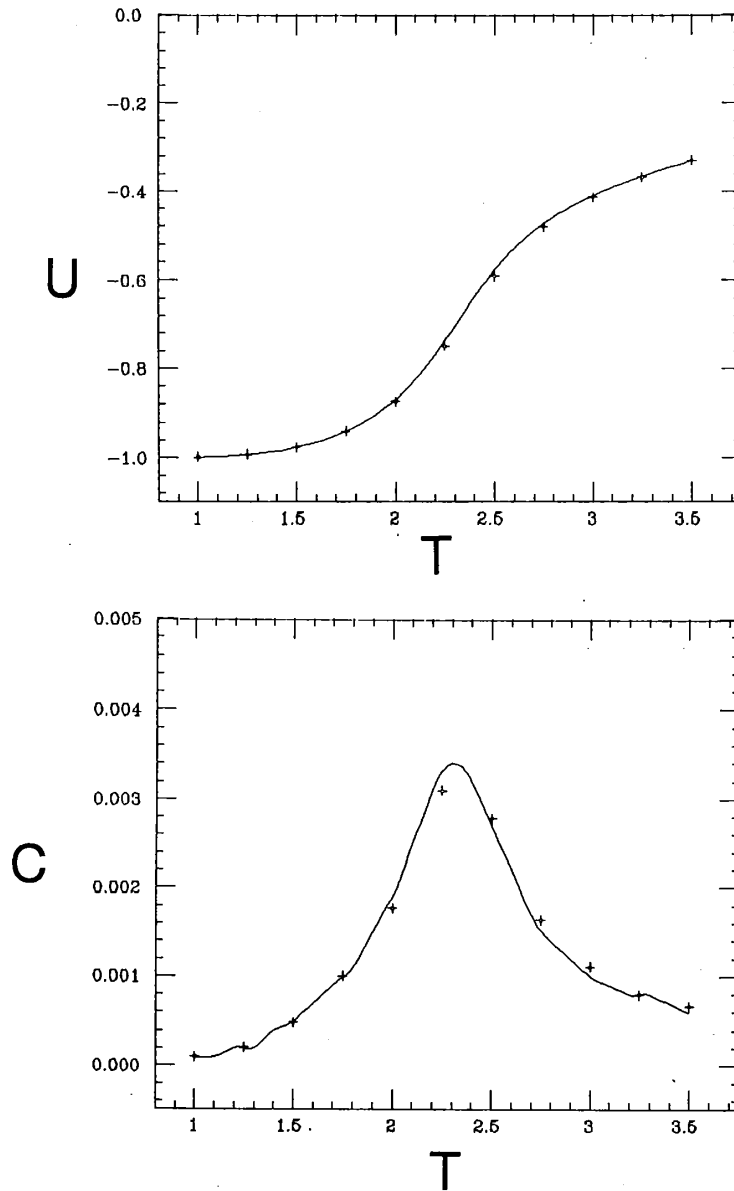


図5. マルチカノニカル法の production run から得られた Ising モデルの内部エネルギー U (a) と比熱 C (b)。

図6には同じ一回の production run から得られた磁化 M の確率分布をプロットした。ここでは、コントラストをはっきりさせるため、それぞれの温度に於けるヒストグラムの最大値を1に規格化した。この図の驚くべき所は、低温に於けるカノニカルモンテカルロシミュレーションでは、磁化の値が+1 または-1 のいずれかの相に一度片寄ってしまうと、エネルギーの高いトンネリング障壁のため、二度とそれから出れなくなってしまうのに対し、マルチカノニカルシミュレーションでは、高温相の状態を通じて、低温の両相が同じ確率で実現されることである。

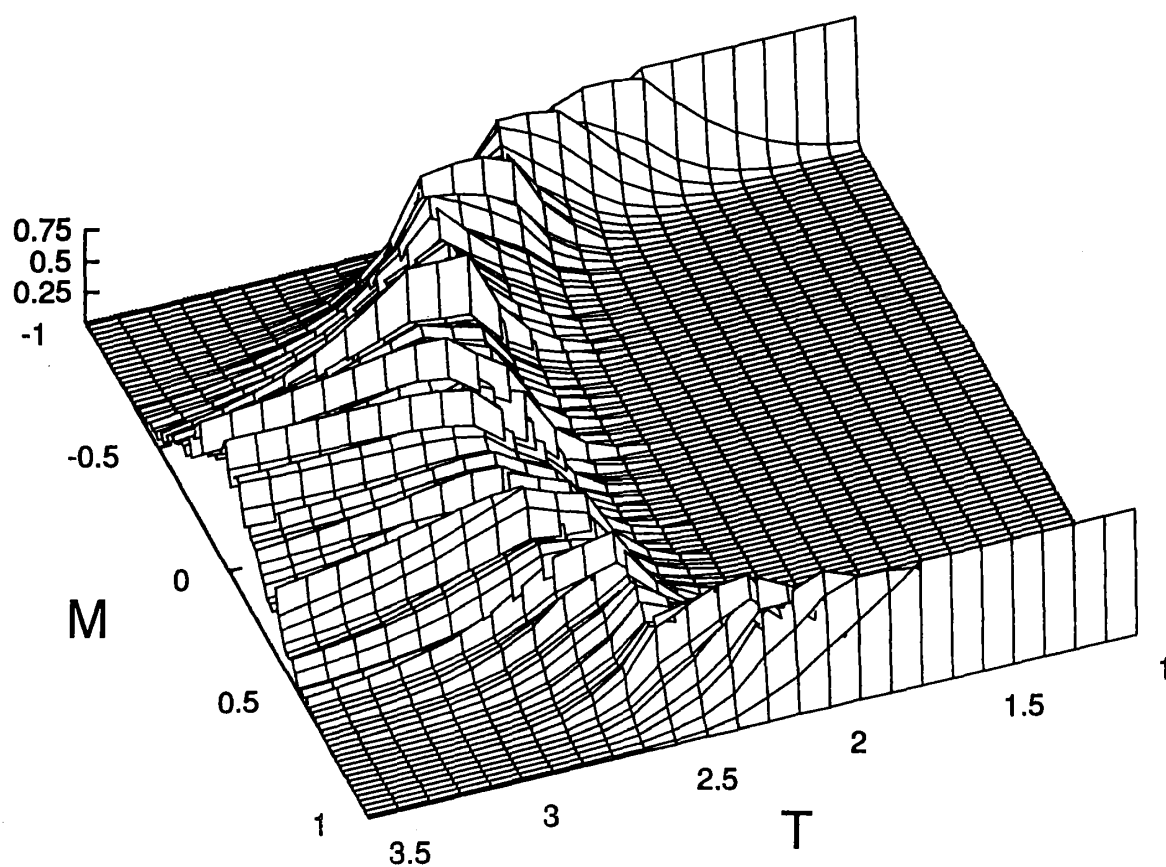


図6. マルチカノニカル法の production run から得られた Ising モデルの磁化 M の確率分布。

話を本題に戻そう。マルチカノニカル法の蛋白質の *ab initio* 立体構造予測への適用は昨年提唱された。²⁶⁾ そこでは、アルゴリズムの有効性を示すために、5個のアミノ酸からなるペプチドであるエンケファリンに於いて、モンテカルロステップ数 10^5 回の production run を一回実行した。その run では、エネルギー空間上のランダムウォークを通じて、以前徐冷モンテカルロ法で求められた最小エネルギー状態²⁸⁾に6回到達していることが判った。²⁶⁾更に、アラニン (Ala)、バリン (Val)、及びグリシン (Gly) がそれぞれ10個つながった、ホモポリマー、 $(\text{Ala})_{10}$ 、 $(\text{Val})_{10}$ 、及び $(\text{Gly})_{10}$ に於けるマルチカノニカルシミュレーションでは、それぞれ一回ずつの production run から、徐冷モンテカルロ法で得られた α -ヘリックス形成傾向性²⁴⁾を再現すると共に、新たに任意の温度に於けるヘリックス-コイル転移の Gibbs 自由エネルギー差、エンタルピー差、及びエントロピー差の計算に成功した。²⁹⁾これらの熱力学量を表2にまとめた。この表から明らかに、室温に於て、アラニンはヘリックス状態を好むが、他のアミノ酸はコイル状態を好む。これらの結果は、最近の小ペプチドに於ける実験結果³⁰⁾と良い一致を示している。

表2. マルチカノニカル法によって得られた、ヘリックス-コイル転移の Gibbs 自由エネルギー差、エンタルピー差、及びエントロピー差 (kcal/mol)。

Peptide	$(\text{Ala})_{10}$			$(\text{Val})_{10}$			$(\text{Gly})_{10}$		
T	250	300	350	250	300	350	250	300	350
ΔG	-4.3	-3.0	-1.9	0.41	0.79	1.1	2.7	3.1	3.6
ΔH	-10.6	-10.1	-10.1	-2.1	-0.94	-1.8	1.3	0.28	-0.29
$T\Delta S$	-6.5	-7.2	-8.3	-1.0	-1.6	-3.8	-0.4	-1.9	-4.1

5. おわりに

本稿ではスーパーコンピュータを使用した物理学の研究例を幾つか紹介した。全然違った分野の研究を取り上げたが、勿論、筆者はいろいろなことを広く浅くやることに意味があると言いたい訳ではない。むしろ、ひとつのことを深く掘り下げる方が大切だと思っている。しかし、計算機という同じ道具を使う以上、学際的な環境の下で、全然別の研究分野で開発された手法が大いに役立つことがあることは忘れてはいけないと思う。特に、徐冷モンテカルロ法やマルチカノニカル法が更に多くの分野で使われるようになることを期待しながら筆を置くことにする。

本研究はコーネル大学物理学科の木下東一郎氏、Rudjer Boskovic 研究所 (Croatia) の Bene Nizic 氏、高エネルギー物理学研究所の川合光氏、ブラウン大学物理学科の Humphrey Maris 氏、日本チバガイギーの菊地武司氏、奈良女子大学化学科の中沢隆氏、京都大学基礎物理学研究所の福来正孝氏、及びフロリダ州立大学の Ulrich Hansmann 氏との共同研究である。コンピュータは奈良女子大学情報処理センター、高エネルギー物理学研究所、分子科学研究所、Cornell National Supercomputer Facility、Stanford Linear Accelerator Center、及び Supercomputer Computations Research Institute のものを使用した。

参考文献

1. Kinoshita, T. (ed.) *Quantum Electrodynamics* (World Scientific, Singapore, 1990).
2. Schwinger, J. *Phys. Rev.* **73**, 416 (1948).
3. Kinoshita, T. and Lindquist, W. B. *Phys. Rev.* **D42**, 636 (1990).
4. Cage, M. E. et al. *IEEE Trans. Instrum. Meas.* **38**, 284 (1989).
5. Van Dyck, R. S., Jr., Schwinberg, P. B. and Dehmelt, H. G. *Phys. Rev. Lett.* **59** 26 (1987).
6. Kinoshita, T., Nizić, B. and Okamoto, Y. *Phys. Rev.* **D31**, 2108 (1985).
7. Jackiw, R. and Weinberg, S. *Phys. Rev.* **D5**, 2396 (1972); Bars, I. and Yoshimura, M. *Phys. Rev.* **D6**, 374 (1972); Fujikawa, K., Lee, B. W. and Sanda, A. I. *Phys. Rev.* **D6**, 2923 (1972); Bardeen, W. A., Gastmans, R. and Lautrup, B. E. *Nucl. Phys.* **B46**, 319 (1972).
8. Kinoshita, T., Nizić, B. and Okamoto, Y. *Phys. Rev.* **D41**, 593 (1990).
9. Bailey, J. et al. *Phys. Lett.* **68B**, 191 (1977); Farley, F. J. M. and Picasso, E. *Annu. Rev. Nucl. Sci.* **29**, 243 (1979).
10. Kinoshita, T., Kawai, H. and Okamoto, Y. *Phys. Lett.* **B254**, 235 (1991); Kawai, H., Kinoshita, T. and Okamoto, Y. *Phys. Lett.* **B260**, 193 (1991).

11. Lanczos, C. *J. Res. Nat. Bur. Stand.* **45**, 255 (1950).
12. Okamoto, Y. and Maris, H.J. *Comput. Phys. Commun.* **76**, 191 (1993).
13. Williams, M.L. and Maris, H.J. *Phys. Rev.* **B35**, 4508 (1985).
14. Yakubo, K., Nakayama, T. and Maris, H.J. *J. Phys. Soc. Jpn.* **60**, 3249 (1991).
15. Medeiros, D. and Cabrera, G.G. *Phys. Rev. B* **43**, 3703 (1991).
16. For example, Kirkpatrick, S., Gelatt, C.D.Jr. and Vecchi, M.P. *Science* **220**, 671 (1983).
17. Berg, B.A. and Neuhaus, T. *Phys. Lett.* **B267**, 249 (1991); *Phys. Rev. Lett.* **68**, 9 (1992).
18. Sippl, M.J., Némethy, G. and Scheraga, H.A. *J. Phys. Chem.* **88**, 6231 (1984), and references therein.
19. Wilson, S.R., Cui, W., Moskowitz, J.W. and Schmidt, K.E. *Tetrahedron Lett.* **29**, 4373 (1988); Kawai, H., Kikuchi, T. and Okamoto, Y. *Protein Engineering* **3**, 85 (1989); Wilson, C. and Doniach, S. *Proteins* **6**, 193 (1989).
20. Kawai, H., Okamoto, Y., Fukugita, M., Nakazawa, T. and Kikuchi, T. *Chemistry Letters* **1991**, 213 (1991); Okamoto, Y., Fukugita, M., Nakazawa, T. and Kawai, H. *Protein Engineering* **4**, 639 (1991).
21. Nakazawa, T., Kawai, H., Okamoto, Y. and Fukugita, M. *Protein Engineering* **5**, 495 (1992).
22. Okamoto, Y., Kikuchi, T., Nakazawa, T. and Kawai, H. *Int. J. Peptide Protein Res.* **42**, 300 (1993).
23. Fukugita, M., Nakazawa, T., Kawai, H. and Okamoto, Y. *Chemistry Letters* **1991**, 1279 (1991).
24. Okamoto, Y. *Proteins* (1994), in press.
25. Okamoto, Y. *Biopolymers* (1994), in press.
26. Hansmann, U. and Okamoto, Y. *J. Comp. Chem.* **14**, 1333 (1993).
27. Ferrenberg, A.M. and Swendsen, R.H. *Phys. Rev. Lett.* **61**, 2635 (1988); *ibid.* **63**, 1658(E) (1989), and references given in the erratum.
28. Okamoto, Y., Kikuchi, T. and Kawai, H. *Chemistry Letters* **1992**, 1275 (1992).
29. Okamoto, Y. and Hansmann, U., Nara Women's University preprint NWU-1/94, submitted for publication.
30. For a review, Chakrabartty, A. and Baldwin, R.L. in *Protein Folding: In Vivo and In Vitro* (eds. Cleland, J. and King, J.) 166–177 (ACS Press, Washington, D.C., 1993).